

ÜBERSICHTSARBEIT

Indirekte Vergleiche und Netzwerk-Metaanalysen

Schätzung von Effekten bei fehlenden direkten Vergleichsstudien –
Teil 22 der Serie zur Bewertung wissenschaftlicher Publikationen

Corinna Kiefer, Sibylle Sturtz, Ralf Bender

ZUSAMMENFASSUNG

Hintergrund: Systematische Übersichten fassen die vorhandenen Studien zu einem Thema strukturiert zusammen. Mit Hilfe von Metaanalysen können bei ausreichend homogener Datenlage gepoolte Effektschätzer berechnet werden. Neben den traditionellen Metaanalysen, in denen direkte Vergleichsstudien bezüglich zweier Interventionen zusammengefasst werden, erweisen sich indirekte Vergleiche und Netzwerk-Metaanalysen als immer bedeutsamer.

Methoden: Basierend auf einer selektiven Literaturrecherche werden verschiedene Methoden für indirekte Vergleiche und Netzwerk-Metaanalysen dargestellt und erläutert. Weiterhin folgen die Beschreibung der wesentlichen Annahmen und Anforderungen dieser Methoden sowie eine Checkliste zur Beurteilung publizierter indirekter Vergleiche und Netzwerk-Metaanalysen.

Ergebnisse: Für den Fall, dass es keine direkten Vergleichsstudien der interessierenden Interventionen gibt, ermöglichen indirekte Vergleiche und Netzwerk-Metaanalysen die Schätzung von Effekten sowie die simultane Analyse von Netzwerken mit mehr als zwei Interventionen. Um eine Netzwerk-Metaanalyse oder einen indirekten Vergleich sinnvoll durchzuführen, müssen die Studien- oder Patientencharakteristika vergleichbar und die beobachteten Effekte ausreichend homogen sein. Ferner soll zwischen direkter und indirekter Evidenz keine bedeutsame Diskrepanz herrschen. Wenn in Studien lediglich gegen eine dritte Intervention aber nicht gegen die interessierende Vergleichsbehandlung getestet wurde, kann diese dritte Intervention als Brückenkomparator dienen, um die interessierenden Interventionen miteinander zu vergleichen.

Schlussfolgerung: Indirekte Vergleiche und Netzwerk-Metaanalysen stellen eine wichtige Weiterentwicklung traditioneller Metaanalysen dar. Eine transparente und ausführliche Dokumentation ist notwendig, um publizierte Ergebnisse von indirekten Vergleichen und Netzwerk-Metaanalysen adäquat beurteilen zu können.

► Zitierweise

Kiefer C, Sturtz S, Bender R: Indirect comparisons and network meta-analyses: estimation of effects in the absence of head-to-head trials—part 22 of a series on evaluation of scientific publications. *Dtsch Arztebl Int* 2015; 112: 803–8. DOI: 10.3238/arztebl.2015.0803

In der medizinischen Forschung wird häufig die relevante Evidenz zu einer Fragestellung mit Hilfe von Übersichten systematisch und transparent gesammelt, bewertet und zusammengefasst (1). Beim Vergleich von genau zwei Interventionen werden die Studienergebnisse der verfügbaren direkten Vergleichsstudien (häufig randomisierte kontrollierte Studien) quantitativ mittels Methoden der Metaanalyse zusammengefasst (1–3). Dies stößt jedoch an Grenzen, wenn es für einen interessierenden Vergleich von zwei Interventionen gar keine direkten Vergleichsstudien gibt oder wenn mehr als zwei Interventionen gleichzeitig miteinander verglichen werden sollen. So gibt es zum Beispiel für den Vergleich neuerer Antikoagulantien bei Patienten mit Vorhofflimmern zwar Studien, in denen jeweils eines der neueren Arzneimittel mit der bisherigen Standardtherapie (Warfarin) abgeglichen wird, jedoch keine Studie, in denen zwei der neueren Antikoagulantien direkt miteinander verglichen werden (4). Ein weiteres Beispiel liefert die Untersuchung verschiedener Prostaglandine zur Einleitung der Geburt. Hier lassen sich 14 Interventionen miteinander vergleichen. Dies sollte anstelle von zahlreichen getrennten paarweisen Vergleichen günstiger in Form einer gemeinsamen Datenanalyse durchgeführt werden (5). In solchen Fällen werden Verfahren für indirekte Vergleiche beziehungsweise Netzwerk-Metaanalysen benötigt.

Seit 2009 gewinnen indirekte Vergleiche und Netzwerk-Metaanalysen immer mehr an Bedeutung (6). Als Erweiterung des bekannten PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Statements (7, 8) wurde auch eine Leitlinie zur Publikation systematischer Übersichten erstellt, die Netzwerk-Metaanalysen beinhalten (9).

Das Ziel dieses Artikels ist es, die grundlegenden Annahmen und Methoden für indirekte Vergleiche und Netzwerk-Metaanalysen zu beschreiben sowie Erläuterungen zu geben, was bei der Beurteilung entsprechender Publikationen zu beachten ist.

Methoden

Im Folgenden werden zunächst die verschiedenen Bezeichnungen, die grundlegenden statistischen Verfahren und die zugehörigen Annahmen für indirekte Vergleiche und Netzwerk-Metaanalysen erläutert. Hierfür wurde eine selektive Literaturrecherche durchgeführt.

Ressort Medizinische Biometrie, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG), Köln: Kiefer, Dr. rer. nat. Sturtz, Prof. Dr. rer. biol. hum. Bender

KASTEN 1

Adjustierter indirekter Vergleich nach Bucher et al. (13)

Unter Verwendung der beiden direkten Schätzer für die Effekte der Intervention *C* relativ zu *A* ($Effekt_{AC}$) sowie der Intervention *C* relativ zu *B* ($Effekt_{BC}$) kann der Effekt der Intervention *B* relativ zur Intervention *A* indirekt geschätzt werden als:

$$Effekt_{AB} = Effekt_{AC} - Effekt_{BC}$$

Die Varianz des indirekten Schätzers $Effekt_{AB}$ ergibt sich als Summe der Varianzen der direkten Schätzer:

$$Varianz_{AB} = Varianz_{AC} + Varianz_{BC}$$

Das zugehörige zweiseitige 95%-Konfidenzintervall lässt sich damit angeben als:

$$\left[Effekt_{AB} - Z_{0,975} \times \sqrt{Varianz_{AB}}; Effekt_{AB} + Z_{0,975} \times \sqrt{Varianz_{AB}} \right]$$

Hierbei bezeichnet $Z_{0,975}$ das 97,5%-Quantil der Standardnormalverteilung, das gerundet den Wert 1,96 ergibt.

Bezeichnungen

Zurzeit existiert in der Fachliteratur noch keine einheitliche Bezeichnung der verschiedenen Methoden für indirekte Vergleiche (10). Analog zu (11) werden hier unter „Methoden für indirekte Vergleiche“ im weiteren Sinn sowohl Verfahren für einfache indirekte Vergleiche von jeweils zwei Interventionen verstanden als auch Verfahren, in denen mehr als zwei Interventionen verglichen und/oder direkte und indirekte Evidenz kombiniert werden. Für Letztere werden in der Literatur die Bezeichnungen Mixed-treatment-comparison-Metaanalyse, Multiple-treatment-Metaanalyse und Netzwerk-Metaanalyse verwendet (10). Es erscheint sinnvoll, von Netzwerk-Metaanalyse zu sprechen, wenn mehr als zwei Interventionen miteinander verglichen werden sollen. Ein indirekter Vergleich im engeren Sinn bezieht sich auf die Synthese ausschließlich indirekter Evidenz bezüglich zweier Interventionen.

Statistische Verfahren

Es herrscht ein allgemeiner wissenschaftlicher Konsens, dass die Anwendung nichtadjustierter indirekter Vergleiche, in denen naiv Ergebnisse einzelner Studienarme unterschiedlicher Studien ohne Berücksichtigung der Randomisierung gegenübergestellt werden, inadäquat ist (11, 12). Daher werden im Folgenden ausschließlich Verfahren für adjustierte indirekte Vergleiche beschrieben, in denen zum Beispiel die in den Studien geschätzten Effekte weiter analysiert werden.

Liegt für den Vergleich von zwei Interventionen keine Evidenz aus direkten Vergleichsstudien vor, wurden jedoch Studien durchgeführt, in denen die zu untersuchenden Interventionen jeweils mit einem gemeinsamen Komparator (zum Beispiel Placebo) verglichen werden, so stellt der adjustierte indirekte Vergleich nach Bucher et al. (13) ein geeignetes Verfahren dar.

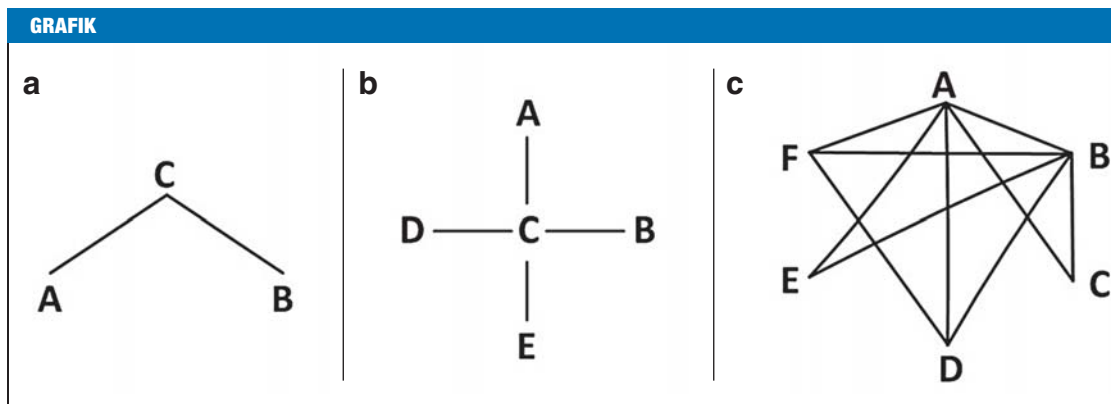
In Teil (a) der *Grafik* ist der einfachste Fall eines Netzwerks mit drei Interventionen A, B und C dargestellt. Durchgezogene Linien stellen dabei direkte Vergleichsstudien dar. Um die beiden Interventionen A und B indirekt zu vergleichen, kann die Evidenz der direkten Vergleichsstudie von A und C und einer zweiten direkten Vergleichsstudie von B und C herangezogen werden. Bei der Intervention C handelt es sich also um den gemeinsamen Komparator der beiden zu vergleichenden Interventionen (Brückenkompator). Der Effekt der Intervention B relativ zu A kann für absolute Effektmaße (zum Beispiel Differenz von Mittelwerten, Risikodifferenz) nach Bucher et al. (13) wie in *Kasten 1* dargestellt indirekt geschätzt werden. Bei relativen Effektmaßen (zum Beispiel Odds Ratio, relatives Risiko) ist zu beachten, dass dieser additive Zusammenhang nur auf der logarithmierten Skala gilt. Die Varianz des indirekten Schätzers ergibt sich dabei aus der Summe der Varianzen der beiden direkten Schätzer. Die Methode nach Bucher et al. (13) kann auch bei sternförmigen Netzwerken (Teil [b] der *Grafik*) angewendet werden, wenn ausschließlich zweiarmige Studien vorliegen. Liegt für einen Vergleich mehr als eine Studie vor, so werden die Studien zunächst metaanalytisch zusammengefasst und der entsprechende Schätzer und dessen Varianz verwendet.

Bei komplexeren Netzwerken (Teil [c] der *Grafik*) wird die Anwendung der Methode nach Bucher et al. (13) immer schwieriger bis unmöglich (10), da dieser weder für die Kombination von Evidenz aus direkten und indirekten Vergleichen noch für den Einbezug mehrarmiger Studien geeignet ist. In diesen Fällen müssen komplexere Modelle im Rahmen von Netzwerk-Metaanalysen herangezogen werden. Netzwerk-Metaanalysen liefern Effektschätzungen für alle möglichen paarweisen Vergleiche innerhalb des Netzwerks. Hierzu werden simultan für jeden paarweisen Vergleich die vorhandene direkte und indirekte Evidenz kombiniert. Die Datenanalyse kann sowohl mit frequentistischen als auch mit Bayes'schen Methoden erfolgen (14). Je nach gewähltem Vorgehen sind verschiedene Aspekte von besonderer Bedeutung. Im Fall von Bayes'schen Methoden spielt besonders die Wahl der Vorinformation in Form von A-priori-Verteilungen eine zentrale Rolle.

Zentrale Annahmen

Allen indirekten Vergleichen liegen drei zentrale Annahmen zugrunde, von denen die beiden ersten auch bei paarweisen Metaanalysen gelten.

- Erstens sollen alle eingeschlossenen Studien hinsichtlich möglicher Effektmodifikatoren (zum Beispiel Studien- oder Patientencharakteristika) vergleichbar sein (Ähnlichkeitsannahme).
- Zweitens soll keine bedeutsame Heterogenität zwischen den Studienergebnissen in den paarweisen Vergleichen vorliegen (Homogenitätsannahme).
- Drittens soll zwischen direkter und indirekter Evidenz keine bedeutsame Diskrepanz (Inkonsistenz) vorliegen (Konsistenzannahme).



Beispiele für Netzwerk-Grafiken

- a) Einfacher indirekter Vergleich
- b) Sternförmiges Netzwerk
- c) Komplexeres Netzwerk mit 6 Interventionen

Beurteilung publizierter indirekter Vergleiche und Netzwerk-Metaanalysen

Im Folgenden wird beschrieben, welche zentralen Aspekte bei der Beurteilung publizierter indirekter Vergleiche und Netzwerk-Metaanalysen zu beachten sind. Diese decken sich notwendigerweise zum Teil mit den Kriterien zur Beurteilung systematischer Übersichten mit traditionellen Metaanalysen (3). Die Anwendung der resultierenden Checkliste (Kasten 2) ermöglicht dem Leser die Beurteilung, ob ein publizierter indirekter Vergleich beziehungsweise eine Netzwerk-Metaanalyse gewisse Kriterien erfüllt. Auf technische Details insbesondere zur statistischen Methodik wird im Rahmen dieser Arbeit verzichtet. Der interessierte Leser wird hierfür auf vorhandene Literatur verwiesen (10, 14–18).

1. Wurden die Fragestellungen a priori festgelegt?

Wie in systematischen Übersichten üblich sollen die zugrundeliegenden Fragestellungen und deren Umsetzung in statistische Hypothesen klar definiert und a priori in einem Studienprotokoll schriftlich festgehalten sein. Abweichungen von der ursprünglichen Planung müssen adäquat beschrieben und begründet werden.

2. Wurde die Durchführung indirekter Vergleiche ausreichend begründet?

Die Anwendung von Methoden für indirekte Vergleiche liefert in der Regel eine geringere Aussagesicherheit als entsprechende direkte Evidenzsynthesen (11). Aus diesem Grund soll immer adäquat begründet werden, wieso für die vorliegende Fragestellung auf Verfahren für indirekte Vergleiche zurückgegriffen wird.

3. Wurde die Wahl der Brückenkomparatoren ausreichend begründet?

Das Ergebnis indirekter Vergleiche hängt unter anderem maßgeblich von der Auswahl der Brückenkomparatoren ab. Liegen mehrere Möglichkeiten vor, können die Ergebnisse unter Umständen gezielt in eine gewünschte Richtung gelenkt werden. Daher soll eine adäquate Begründung für die Wahl der jeweiligen Brückenkomparatoren angegeben werden.

4. Wurde eine systematische und vollständige Literaturrecherche durchgeführt und detailliert beschrieben?

Wie bei systematischen Übersichten mit traditionellen Metaanalysen ist eine systematische und vollständige Literaturrecherche notwendig (7, 8). Die Vollständigkeit der relevanten Literatur muss hier nicht nur in Bezug auf die im Fokus stehenden Interventionen gelten, sondern auch für die Brückenkomparatoren.

5. Wurden a priori festgelegte Ein- und Ausschlusskriterien für die Studien angewendet und klar beschrieben?

Wie in systematischen Übersichten üblich wird anhand der Ein- und Ausschlusskriterien darüber entschieden, welche der in der Literaturrecherche gefundenen Studien eingeschlossen werden. Die Identifizierung der relevanten Literatur soll idealerweise in Form eines Flowcharts transparent dargestellt werden (7, 8).

6. Lag eine vollständige Berichterstattung aller relevanten Daten vor?

Eine vollständige Berichterstattung ist notwendig zur Vermeidung verzerrter Ergebnisse. Dies betrifft hier die vollständige Darstellung der wesentlichen Studiencharakteristika, der Studienbewertungen, der einzelnen Studienergebnisse (Effektschätzungen und Konfidenzintervalle) sowie der gepoolten Effektschätzungen inklusive Konfidenzintervallen für alle relevanten Endpunkte und Vergleiche sowie alle relevanten Subgruppen. Zumindest bei komplexeren Netzwerken soll außerdem eine Grafik des betrachteten Netzwerks mit einer Beschreibung der Netzwerk-Geometrie, das heißt der wesentlichen Eigenschaften des Netzwerks, präsentiert werden (9, 19, 20).

7. Wurden die zentralen Annahmen Ähnlichkeit, Homogenität und Konsistenz untersucht und wurde adäquat mit den Ergebnissen daraus umgegangen?

Wie in paarweisen Metaanalysen soll die Ähnlichkeit der eingeschlossenen Studien anhand ihrer zentralen Charakteristika untersucht werden; dies gilt aber jetzt über alle betrachteten Interventionen hinweg. Eine zen-

KASTEN 2

Checkliste zur Beurteilung indirekter Vergleiche bzw. Netzwerk-Metaanalysen

1. Wurden die Fragestellungen a priori festgelegt?
 - klare Beschreibung der Fragestellung
 - Umsetzung in statistische Hypothesen
 - Begründung bei Abweichungen von der ursprünglichen Planung
2. Wurde die Durchführung indirekter Vergleiche ausreichend begründet?
3. Wurde die Wahl der Brückenkomparatoren ausreichend begründet?
4. Wurde eine systematische und vollständige Literaturrecherche durchgeführt und detailliert beschrieben?
 - Für die primär interessierenden Interventionen?
 - Für die Brückenkomparatoren?
5. Wurden a priori festgelegte Ein- und Ausschlusskriterien für die Studien angewendet und klar beschrieben?
6. Lag eine vollständige Berichterstattung aller relevanten Daten vor?
 - Charakteristika aller eingeschlossenen Studien
 - Bewertungen aller eingeschlossenen Studien
 - Grafik des Netzwerks, Beschreibung der Netzwerk-Geometrie
 - Für alle relevanten Endpunkte, Vergleiche und Subgruppen:
 - Einzelergebnisse aller Studien (Effektschätzer und Konfidenzintervalle)
 - Effektschätzer und Konfidenzintervalle aus paarweisen Metaanalysen
7. Wurden die zentralen Annahmen untersucht und wurde adäquat mit den Ergebnissen daraus umgegangen?
 - Ähnlichkeit
 - Homogenität
 - Konsistenz
8. Wurden adäquate statistische Verfahren angewendet und detailliert beschrieben?
 - Anwendung adjustierter indirekter Vergleiche
 - Umgang mit mehrarmigen Studien
 - Modelle mit zufälligen oder festen Effekten
 - technische Details (insbesondere bei Bayes'schen Verfahren)
 - Programmcode
 - Sensitivitätsanalysen
9. Wurden die Limitationen ausreichend beschrieben und diskutiert?
 - Qualität und Vollständigkeit der Datenbasis
 - methodische Unklarheiten, Sensitivitätsanalysen
 - Verletzungen der zentralen Annahmen

trale Rolle spielt hierbei das bekannte PICOS-Schema, das heißt die Betrachtung der Population (P), der Intervention (I), des Komparators (C für Comparator), der Endpunkte (O für Outcome) und des Studiendesigns (S) (7, 8). Wichtige Informationen können Vergleiche der Studien hinsichtlich relevanter Patientencharakteristika sowie der Studienarme einer sinnvollen Referenzintervention hinsichtlich relevanter Endpunkte liefern.

Analog zu paarweisen Metaanalysen soll die Homogenität mit üblichen Verfahren wie Forest Plots und Heterogenitätsmaßen untersucht werden, wobei die Entscheidungskriterien vorab definiert sein sollen (7, 8). Je nach Größe des Netzwerks kann dies allerdings sehr umfangreich sein, da alle vorhandenen Kombinationen von jeweils zwei Interventionen betrachtet werden müssen.

Die Konsistenz kann in der Regel nur bei gleichzeitigem Vorliegen von direkter und indirekter Evidenz für einen Vergleich untersucht werden. Das bedeutet, dass einfache indirekte Vergleiche, die über genau einen Brückenkomparator verbunden sind, immer mit erhöhter Unsicherheit behaftet sind, da eine der zentralen Annahmen nicht untersucht werden kann. Zur Untersuchung der Konsistenz gibt es eine Reihe von Verfahren (16, 21–24). Zur Beurteilung, ob diese Verfahren adäquat angewendet wurden, ist die Unterstützung durch einen erfahrenen Biometriker notwendig.

Mit den Ergebnissen der Überprüfung auf Ähnlichkeit, Homogenität und Konsistenz soll adäquat umgegangen werden. Das kann auch bedeuten, dass aufgrund einer starken Verletzung dieser Annahmen ein indirekter Vergleich nicht sinnvoll möglich ist (23). Unter Umständen kann durch eine Änderung oder eine Aufspaltung des Netzwerks in Subgruppen oder die Aufnahme relevanter Kovariablen ein sinnvoller indirekter Vergleich ermöglicht werden (10, 25).

8. Wurden adäquate statistische Verfahren angewendet und detailliert beschrieben?

Aufgrund der Komplexität einiger Verfahren für Netzwerk-Metaanalysen ist deren Beurteilung eine besondere Herausforderung, die nur von einem erfahrenen Biometriker geleistet werden kann. Erforderlich sind hierfür eine sehr ausführliche Beschreibung der angewendeten Methoden (9) und häufig auch die Darstellung des Programmcodes. Eine Hilfestellung liefern publizierte Checklisten (26–28). Zentrale Aspekte hierbei sind:

- ob adjustierte indirekte Vergleiche verwendet wurden
- ob angemessene Modelle mit festen oder zufälligen Effekten gewählt wurden
- ob mit mehrarmigen Studien adäquat umgegangen wurde
- ob unklare Aspekte mit Hilfe von Sensitivitätsanalysen untersucht wurden.

In Abhängigkeit davon, ob Bayes'sche oder frequentistische Verfahren angewendet wurden, spielen weitere technische Details eine Rolle, die hier nicht näher beschrieben werden können.

9. Wurden die Limitationen ausreichend beschrieben und diskutiert?

Wie in jeder wissenschaftlichen Untersuchung sollen mögliche Limitationen, welche die Ergebnissicherheit einschränken, ausreichend berichtet und diskutiert werden. Bei indirekten Vergleichen und Netzwerk-Metaanalysen spielen insbesondere die Qualität und Vollständigkeit der Datenbasis, methodische Unklarheiten und damit verbundene Sensitivitätsanalysen sowie mögliche Verletzungen der zentralen Annahmen eine Rolle. Eine entsprechende Beurteilung ist von einem erfahrenen Biometriker durchzuführen.

Ergebnisse

Die beschriebene Checkliste (*Kasten 2*) wurde beispielhaft auf publizierte adjustierte indirekte Vergleiche sowie eine publizierte Netzwerk-Metaanalyse angewendet. Die Ergebnisse sind in der *Tabelle* dargestellt. Baker und Phung (4) untersuchten anhand adjustierter indirekter Vergleiche neuere Antikoagulantien bei Patienten mit Vorhofflimmern. Sie schlossen insgesamt vier randomisierte kontrollierte Studien zum Vergleich von Apixaban (1 Studie), Dabigatran (2 Studien) und Rivaroxaban (1 Studie) jeweils mit der bisherigen Standardtherapie Warfarin ein, wobei Warfarin als Brückenkomparator für die indirekten Vergleiche gewählt wurde. Eine Bayes'sche Netzwerk-Metaanalyse wurde von Alfirevic et al. (5) zum Vergleich des Nutzens und Schadens verschiedener Prostaglandine zur Einleitung der Geburt durchgeführt. Je nach verfügbarer Datenlage für den jeweiligen Endpunkt wurden Netzwerke mit bis zu 14 Interventionen und einer Datenbasis von insgesamt 280 Studien ausgewertet.

Beide Publikationen sind Beispiele für eine im Wesentlichen sorgfältige Durchführung und Berichterstattung indirekter Vergleiche. Teilweise wäre allerdings die Angabe von mehr Details wünschenswert gewesen. Bei Anwendung der Checkliste (*Kasten 2*) ergeben sich zudem die folgenden Limitationen.

Baker und Phung (4) begründeten die Wahl von Warfarin als Brückenkomparator nicht ausreichend. Sie geben zwar an, dass Warfarin die bisherige Standardtherapie sei, für die zumindest für alle drei eingeschlossenen Substanzen bereits die Nicht-Unterlegenheit gezeigt wurde, es bleibt jedoch unklar, ob diese Wahl möglicherweise selektiv ist oder es tatsächlich die einzige Substanz ist, zu der direkt vergleichende Studien zu allen drei interessierenden Substanzen vorliegen. So wäre ein Placebo als weiterer potenzieller Brückenkomparator denkbar, dieser wird allerdings von den Autoren nicht weiter diskutiert.

Beide Publikationen beschreiben eine detaillierte Informationsbeschaffung. Allerdings fehlt bei Baker und Phung (4) eine Liste der ausgeschlossenen Referenzen. Alfirevic et al. (5) stützen sich auf nur eine Sekundärquelle (Cochrane Pregnancy and Childbirth Group Register) mit unklarer Aktualisierung. Es wird nicht nachvollziehbar dargestellt, wie diese Sekundärquelle durchsucht wurde. Dadurch bleibt es im Ungewissen, ob der Studienpool aktuell und vollständig ist.

In beiden Publikationen fehlt die Untersuchung der Homogenitätsannahme in den paarweisen Metaanalysen. Alfirevic et al. (5) untersuchten diese lediglich im Rahmen der Netzwerk-Metaanalyse für alle paarweisen Vergleiche gemeinsam; dieses Vorgehen ist jedoch nicht ausreichend. Bei Baker und Phung (4) fehlt zudem die Darstellung der gepoolten Ergebnisse des Dabigatran-Warfarin-Vergleichs. Aufgrund der fehlenden Untersuchung der Homogenitätsannahme ist es möglich, dass bei beiden Publikationen paarweise Vergleiche mit bedeutsamer Heterogenität zwischen den Studienergebnissen eingehen.

TABELLE

Beurteilung der indirekten Vergleiche und der Netzwerk-Metaanalyse in den Datenbeispielen

| Item der Checkliste | Baker und Phung (4) | Alfirevic et al. (5) |
|---|---------------------|----------------------|
| 1. Fragestellungen a priori festgelegt? | ✓ | ✓ |
| 2. Durchführung indirekter Vergleiche begründet? | ✓ | ✓ |
| 3. Wahl der Brückenkomparatoren begründet? | — ^{*1} | ✓ |
| 4. Literaturrecherche vollständig? | (✓) ^{*2} | (✓) ^{*5} |
| 5. Ein- und Ausschlusskriterien klar beschrieben? | ✓ | ✓ |
| 6. Vollständige Berichterstattung aller Daten? | (✓) ^{*3} | ✓ |
| 7. Zentrale Annahmen überprüft? | — ^{*4} | — ^{*4} |
| 8. Adäquate statistische Verfahren angewendet? | ✓ | ✓ |
| 9. Limitationen beschrieben? | ✓ | ✓ |

^{*1} Es ist unklar, ob es sich um eine selektive Auswahl handelt oder nicht.

^{*2} Es fehlt eine Liste der ausgeschlossenen Referenzen.

^{*3} Es fehlt die Darstellung der gepoolten Schätzungen für den paarweisen Vergleich von Dabigatran und Warfarin.

^{*4} Die Untersuchung der Homogenität ist nicht ausreichend.

^{*5} Es wird nicht nachvollziehbar dargestellt, wie die verwendete Sekundärquelle durchsucht wurde, so dass unklar ist, ob der Studienpool aktuell und vollständig ist.

Diskussion

Indirekte Vergleiche und Netzwerk-Metaanalysen ermöglichen im Rahmen systematischer Übersichten die Schätzung von Effekten für den Vergleich von Interventionen, auch wenn es hierzu keine direkten Vergleichsstudien gibt, sowie die simultane Analyse von Netzwerken mit mehr als zwei Interventionen. Durch die Anwendung adjustierter indirekter Vergleiche ist es zum Beispiel möglich, Nutzen und Schaden neuerer Antikoagulantien (Apixaban, Dabigatran, Rivaroxaban) bei Patienten mit Vorhofflimmern miteinander zu vergleichen, obwohl es hierzu keine direkten Vergleichsstudien gibt. Des Weiteren ist es mit Hilfe einer Netzwerk-Metaanalyse machbar, zum Beispiel den Nutzen und Schaden von 14 Interventionen (Prostaglandine, keine Behandlung, Placebo) zur Geburtseinleitung in einer gemeinsamen Auswertung simultan miteinander zu vergleichen. Indirekte Vergleiche und Netzwerk-Metaanalysen können somit Ergebnisse liefern, die zur Entwicklung von Leitlinien und für evidenzbasierte Entscheidungen im Gesundheitswesen grundlegend von Bedeutung sind. Daher stellen sie wichtige Erweiterungen der traditionellen paarweisen Metaanalysen dar. Für eine adäquate Anwendung dieser Methoden werden allerdings starke Annahmen benötigt, die anhand der verfügbaren Daten häufig nicht vollständig überprüfbar sind. Deshalb besitzen Ergebnisse aus indirekten Vergleichen und Netzwerk-Metaanalysen in der Regel eine geringere Aussagesicherheit als Ergebnisse aus adäquaten paarweisen Metaanalysen direkter Vergleichsstudien. Eine transparente und ausführliche Dokumentation ist notwendig, um publizierte Ergebnisse von indirekten Vergleichen und Netzwerk-Metaanalysen adäquat beurteilen zu können. Eine Hilfestellung hierzu bietet die in dieser Arbeit präsentierte einfache Checkliste.

KERNAUSSAGEN

- Indirekte Vergleiche und Netzwerk-Metaanalysen ermöglichen im Rahmen systematischer Übersichten die Schätzung von Effekten, wenn es keine direkten Vergleichsstudien gibt oder mehrere Interventionen gleichzeitig miteinander verglichen werden sollen.
- Die zentralen Annahmen der Methoden für indirekte Vergleiche und Netzwerk-Metaanalysen sind ausreichende Ähnlichkeit, Homogenität und Konsistenz.
- Nichtadjustierte indirekte Vergleiche, das heißt naive Vergleiche einzelner Studienarme verschiedener Studien, stellen keine adäquate Analyseverfahren dar.
- Eine transparente und ausführliche Dokumentation ist notwendig, um publizierte Ergebnisse von indirekten Vergleichen und Netzwerk-Metaanalysen adäquat beurteilen zu können.
- Zur Beurteilung der statistischen Methoden für indirekte Vergleiche und Netzwerk-Metaanalysen ist ein erfahrener Biometriker erforderlich.

Danksagung

Die Autoren möchten sich bei Elke Hausner für die Bewertung der Literaturrecherchen in den Datenbeispielen sowie bei Ulrich Grouven für die hilfreichen redaktionellen Hinweise bedanken.

Interessenkonflikt

Die Autoren erklären, dass kein Interessenkonflikt besteht.

Manuskriptdaten

eingereicht: 16. 7. 2015, revidierte Fassung angenommen: 24. 8. 2015

LITERATUR

1. Egger M, Davey Smith G, Altman DG (eds.): Systematic reviews in health care: meta-analysis in context. London: BMJ Books 2001.
2. Ziegler A, Lange S, Bender R: Systematische Übersichten und Meta-Analysen. Dtsch Med Wochenschr 2007; 132 (Suppl. 1): e48–52.
3. Ressing M, Blettner M, Klug SJ: Systematic literature reviews and meta-analyses: Part 6 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2009; 106: 456–63.
4. Baker WL, Phung OJ: Systematic review and adjusted indirect comparison meta-analysis of oral anticoagulants in atrial fibrillation. Circ Cardiovasc Qual Outcomes 2012; 5: 711–9.
5. Alfirevic Z, Keeney E, Dowswell T, et al.: Labour induction with prostaglandins: A systematic review and network meta-analysis. BMJ 2015; 350: h217.
6. Lee AW: Review of mixed treatment comparisons in published systematic reviews shows marked increase since 2009. J Clin Epidemiol 2014; 67: 138–43.
7. Moher D, Liberati A, Tetzlaff J, Altman DG: Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. BMJ 2009; 339: b2535.
8. Liberati A, Altman DG, Tetzlaff J, et al.: The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009; 339: b2700.
9. Hutton B, Salanti G, Caldwell DM, et al.: The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: Checklist and explanations. Ann Intern Med 2015; 162: 777–84.
10. Salanti G: Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: Many names, many benefits, many concerns for the next generation evidence synthesis tool. Res Syn Methods 2012; 3: 80–97.
11. Bender R, Schwenke C, Schmoor C, Hauschke D: Stellenwert von Ergebnissen aus indirekten Vergleichen – Gemeinsame Stellungnahme von IQWiG, GMDS und IBS-DR. Köln, 02. Februar 2012. www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWiG_GMDS_IBS_DR.pdf (last accessed on 21 September 2015).

12. Higgins JPT, Deeks JJ, Altman DG, on behalf of the Cochrane Statistical Methods Group (eds.): Special topics in statistics. In: Higgins JPT, Deeks JJ, Altman DG, Green S (eds.): Cochrane Handbook for Systematic Reviews of Interventions. Chichester: Wiley 2008; 481–529.
13. Bucher HC, Guyatt GH, Griffith LE, Walter SD: The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997; 50: 683–91.
14. Hoaglin DC, Hawkins N, Jansen JP, et al.: Conducting indirect-treatment-comparison and network-meta-analysis studies: Report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: Part 2. Value Health 2011; 14: 429–37.
15. Dias S, Sutton AJ, Ades AE, Welton NJ: Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. Med Decis Making 2013; 33: 607–17.
16. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE: Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. Med Decis Making 2013; 33: 641–56.
17. Rücker G: Network meta-analysis, electrical networks and graph theory. Res Syn Methods 2012; 3: 312–24.
18. Salanti G, Higgins JPT, Ades A, Ioannidis JPA: Evaluation of networks of randomized trials. Stat Methods Med Res 2007; 17: 279–30.
19. Salanti G, Kavvoura FK, Ioannidis JPA: Exploring the geometry of treatment networks. Ann Intern Med 2008; 148: 544–53.
20. Rücker G, Schwarzer G: Automated drawing of network plots in network meta-analysis. Res Syn Methods 2015 Jun 9. DOI: 10.1002/rjsm.1143 (Epub ahead of print).
21. Dias S, Welton NJ, Caldwell DM, Ades AE: Checking consistency in mixed treatment comparison meta-analysis. Stat Med 2010; 29: 932–44.
22. Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR: Consistency and inconsistency in network meta-analysis: Concepts and models for multi-arm studies. Res Syn Methods 2012; 3: 98–110.
23. Donegan S, Williamson P, D'Alessandro U, Tudur Smith C: Assessing key assumptions of network meta-analysis: A review of methods. Res Syn Methods 2013; 4: 291–323.
24. Krahn U, Binder H, König J: A graphical tool for locating inconsistency in network meta-analyses. BMC Med Res Methodol 2013; 13: 35.
25. Sturtz S, Bender R: Unsolved issues of mixed treatment comparison meta-analysis: Network size and inconsistency. Res Syn Methods 2012; 3: 300–11.
26. Jansen JP, Fleurence R, Devine B, et al.: Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: Report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: Part 1. Value Health 2011; 14: 417–28.
27. Jansen JP, Trikalinos T, Cappelleri JC, et al.: Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: An ISPOR-AMCP-NPC Good Practice Task Force Report. Value Health 2014; 17: 157–73.
28. Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ, Dias S: Evidence synthesis for decision making 7: A reviewer's checklist. Med Decis Making 2013; 33: 679–91.

Anschrift für die Verfasser

Prof. Dr. rer. biol. hum. Ralf Bender
 Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)
 Im Mediapark 8
 50670 Köln
 Ralf.Bender@iqwig.de

Zitierweise

Kiefer C, Sturtz S, Bender R: Indirect comparisons and network meta-analyses: estimation of effects in the absence of head-to-head trials—part 22 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2015; 112: 803–8.
 DOI: 10.3238/arztebl.2015.0803

 The English version of this article is available online:
www.aerzteblatt-international.de